Problem 1

a)

E-step:

1. Set $q(c) = p(c|\pi, x, \theta)$

$$p(c|\pi, x, \theta) = \prod_{i=1}^{n} p(c_i|\theta, x_i, \pi)$$

$$p(c_i = k|\theta, x_i, \pi) = \frac{p(x_i|c_i = k, \theta)p(c_i = k|\pi)}{\sum_{j=1}^{K} p(x_i|c_i = j, \theta)p(c_i = j|\pi)}$$

$$= \frac{\pi_k Binomial(x_i, 20, \theta_k)}{\sum_{j=1}^{K} \pi_j Binomial(x_i, 20, \theta_j)} = \phi_i(k)$$

2. Calculate $\sum_c q(c) lnp(x, c|\pi, \theta)$

$$\sum_c q(c) lnp(x, c|\pi, \theta) = \sum_{i=1}^{n}\sum_{k=1}^{K} q(c_i = k)(lnp(x_i|\theta, c_i = k) + ln\pi_k)$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{K} \phi_i(k)\left(ln\theta_k^{x_i}(1-\theta_k)^{20-x_i} + ln\pi_k\right) + const$$

M-step:

1. $\nabla_{\theta_k} L = 0$

$$\nabla_{\theta_k} L = \sum_{i=1}^{n} \phi_i(k)\left(\frac{x_i}{\theta_k} - \frac{20 - x_i}{1 - \theta_k}\right) = 0$$

$$\theta_k = \frac{\sum_{i=1}^{n} \phi_i(k) x_i}{20 n_k}, \qquad n_k = \sum_{i=1}^{n} \phi_i(k)$$

2. $\nabla_{\pi_k} L = 0$

$$\pi_k = \frac{n_k}{n}$$

1. Initializing, $\boldsymbol{\pi}^0$, $\boldsymbol{\theta}^0$ in some way.

2. For interation t = 1,2, ⋯, T

   (a) E-step: for i = 1,2,⋯,n and k = 1,2,⋯,K , set:

$$\phi_i^{(t)}(k) = \frac{\pi_k^{(t-1)} \theta^{(t-1)}{}_k^{x_i} \left(1 - \theta_k^{(t-1)}\right)^{20-x_i}}{\sum_{j=1}^{K} \pi_j^{(t-1)} \theta^{(t-1)}{}_j^{x_i} \left(1 - \theta_j^{(t-1)}\right)^{20-x_i}}$$

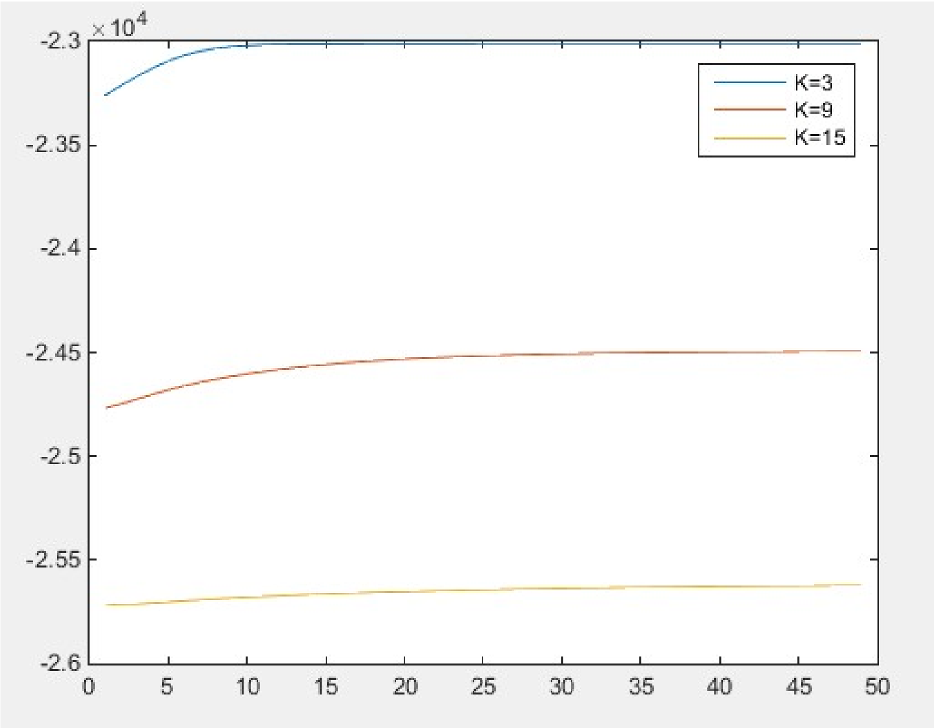
   (b) M-step: set:

$$n_k^{(t)} = \sum_{i=1}^{n} \phi_i^{(t)}(k), \qquad \theta_k^{(t)} = \frac{\sum_{i=1}^{n} \phi_i^{(t)}(k) x_i}{20 n_k^{(t)}}, \qquad \pi_k^{(t)} = \frac{n_k^{(t)}}{n}$$

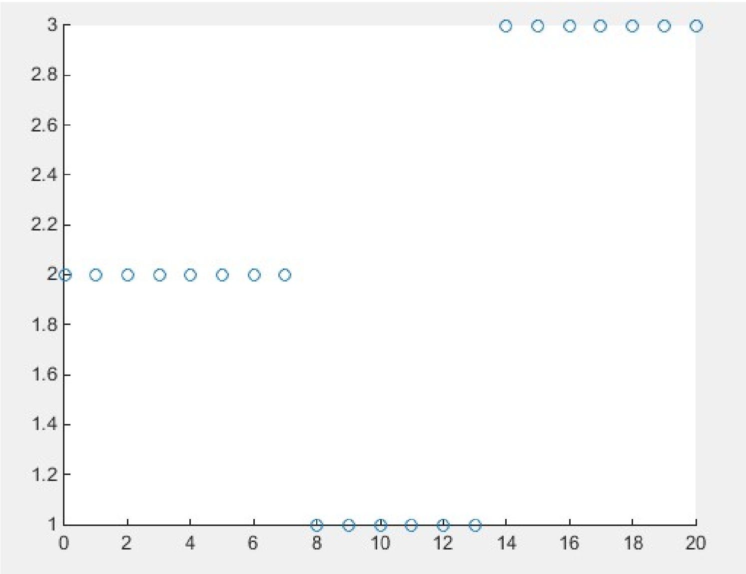
(c)Calculate $lnp(x|\boldsymbol{\pi}^{(t)}, \boldsymbol{\theta}^{(t)})$ to assess convergence

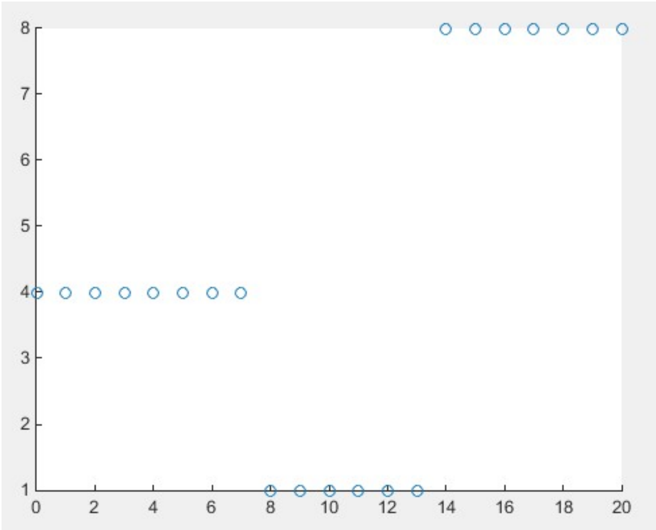
$$lnp(x|\boldsymbol{\pi}^{(t)}, \boldsymbol{\theta}^{(t)}) = \sum_{i=1}^{n}\sum_{k=1}^{K} \phi_i^{(t)}(k)(ln\theta^{(t)}{}_k^{x_i}\left(1-\theta_k^{(t)}\right)^{20-x_i} + ln\pi_k^{(t)})$$

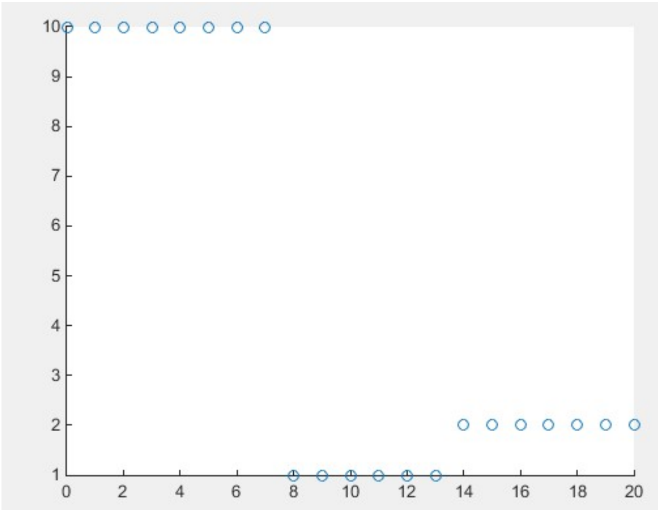
b) Log marginal likehood



c) Use $q(c|x,\theta)$ to find most probable clusters for three model







Problem 2

a)

$$q(c_i = j) \propto e^{E\left[lnp\left(x_i \middle| \theta_j\right) + lnp(c_i = j|\pi)\right]}$$

$$\propto e^{E[ln\frac{20!}{x_i!(20-x_i)!}\theta_j^{x_i}(1-\theta_j)^{20- \quad i} + ln\pi_j]}$$

$$\propto e^{x_i E[ln \quad j] + (20-x_i)E[\ln(1-\theta_j)] + E[ln\pi_j]}$$

$$= \text{Discrete}(\phi_i)$$

$$q(\pi) \propto e^{\sum_{i=1}^{n} E[lnp(c_i|\pi)] + lnp(\pi)}$$

$$\propto e^{\sum_{i=1}^{n}\sum_{j=1}^{K}\phi_i(j)ln\pi_j + \sum_{j=1}^{K}(\alpha-1)ln\pi_j}$$

$$\propto \prod_{j=1}^{K} \pi_j^{\alpha-1+\sum_{i=1}^{n}\phi_i(j)}$$

$$= \text{Dirichlet}(\alpha')$$

$$\alpha_j' = \alpha + n_j\,, \qquad n_j = \sum_{i=1}^{n} \phi_i(j)$$

$$q(\theta_j) \propto e^{\sum_{i=1}^{n} E[lnp(x_i|c_i=j,\theta_j)]+E[lnp(\theta_j)]}$$

$$\propto e^{\sum_{i=1}^{n} \phi_i(j) ln\frac{20!}{x_i!(20-x_i)!}\theta_j^{x_i}(1-\theta_j)^{20-x_i}} \theta_j^{a-1}(1-\theta_j)^{b-1}$$

$$= Beta(a', b')$$

$$a_j' = \sum_{i=1}^{n} \phi_i(j) x_i + a\,, \qquad b_j' = \sum_{i=1}^{n} \phi_i(j)(20 - x_i) + b$$

Input: data $x_1, \ldots, x_n$, number of clusters.

Output: parameters for $q(\theta_j), q(\pi)$ and $q(c_i)$

1. Initialize $\phi_i^{(0)}, \left(\alpha_1^{(0)}, \ldots, \alpha_K^{(0)}\right), (a^{(0)}, b^{(0)})$ in some way.
2. At interation t:
   (a) Update $q(c_i)$ for $i = 1, \ldots n$ by setting:

$$\phi_i^{(t)}(j) = \frac{e^{x_i\left(\psi\left(a_j^{(t-1)}\right)-\psi\left(a_j^{(t-1)}+b_j^{(t-1)}\right)\right)+(20-x_i)\left(\psi\left(b_j^{(t-1)}\right)-\psi\left(a_j^{(t-1)}+b_j^{(t-1)}\right)\right)+\psi\left(\alpha_j'\right)-\psi(\sum_j^K \alpha'_j)}}{\sum_{k=1}^{K} e^{x_i\left(\psi\left(a_k^{(t-1)}\right)-\psi\left(a_k^{(t-1)}+b_k^{(t-1)}\right)\right)+(20-x_i)\left(\psi\left(b_k^{(t-1)}\right)-\psi\left(a_k^{(t-1)}+b_k^{(t-1)}\right)\right)+\psi\left(\alpha_k'\right)-\psi(\sum_j^K \alpha'_j)}}$$

   (b) Set $n_j^{(t)} = \sum_{i=1}^{n} \phi_i(j)\,, for\ j = 1, \ldots, K$

   (c) Update $q(\pi)$ by setting:

$$\alpha_j^{(t)} = \alpha + n_j^{(t)}$$

   $for\ j = 1, \ldots, K$

   (d) Update $q(\theta_j)$ by setting:

$$a_j^{(t)} = \sum_{i=1}^{n} \phi_i^{(t)}(j) x_i + a\,, \qquad b_j^{(t)} = \sum_{i=1}^{n} \phi_i^{(t)}(j)(20 - x_i) + b$$

   (e) Calculate Objective function:

$$L = E_q[lnp(x, c, \theta, \pi)] - E_q[lnq]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{K} \phi_i^{(t)}(j) \left[x_i\left(\psi\left(a_j^{(t)}\right) - \psi\left(a_j^{(t)}+b_j^{(t)}\right)\right) + (20 - x_i)\left(\psi\left(b_j^{(t)}\right) - \psi\left(a_j^{(t)}+b_j^{(t)}\right)\right) + \psi\left(\alpha_j^{(t)}\right) - \psi(\sum_k \alpha_k^{(t)})\right]$$

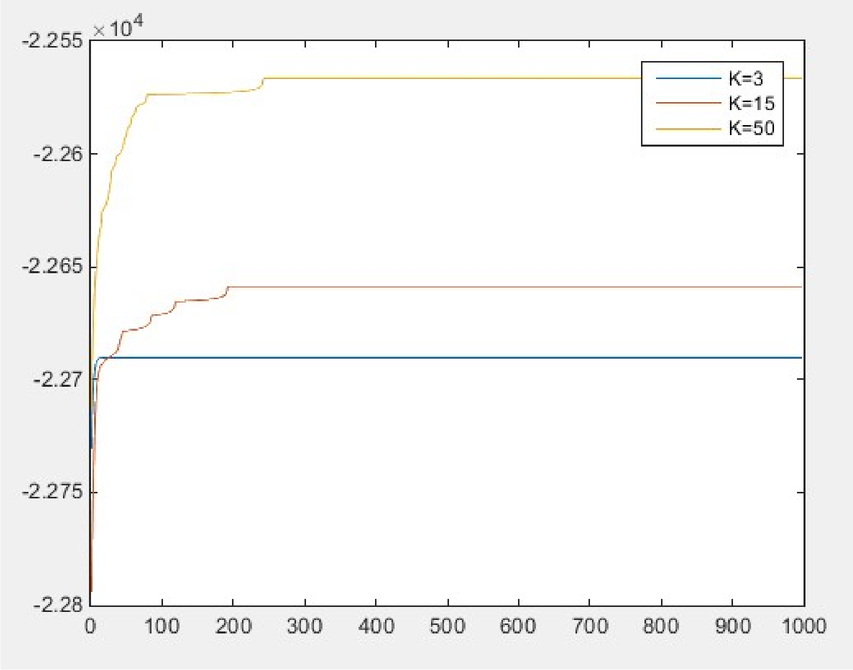
$$+ \sum_{j=1}^{K} (a-1)\left(\psi\left(a_j^{(t)}\right) - \psi\left(a_j^{(t)}+b_j^{(t)}\right)\right) + (b-1)\left(\psi\left(b_j^{(t)}\right) - \psi\left(a_j^{(t)}+b_j^{(t)}\right)\right)$$

$$+\sum_{j=1}^{K}(\alpha-1)(\psi\left(\alpha_j^{(t)}\right)-\psi\left(\sum_k \alpha_k^{(t)}\right))-\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{j=1}^{K}\phi_i^{(t)}(j)ln\phi_i^{(t)}(j)$$

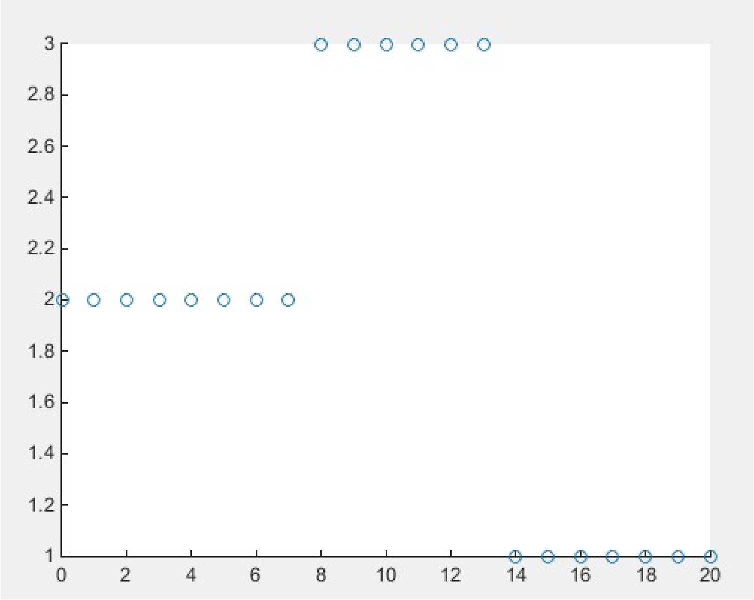
$$+\sum_{j=1}^{K} lnB(a_j^{(t)},b_j^{(t)})-(a_j^{(t)}-1)\psi(a_j^{(t)})-(b_j^{(t)}-1)\psi(b_j^{(t)})+(a_j^{(t)}+b_j^{(t)}-2)\psi(a_j^{(t)}+b_j^{(t)})$$

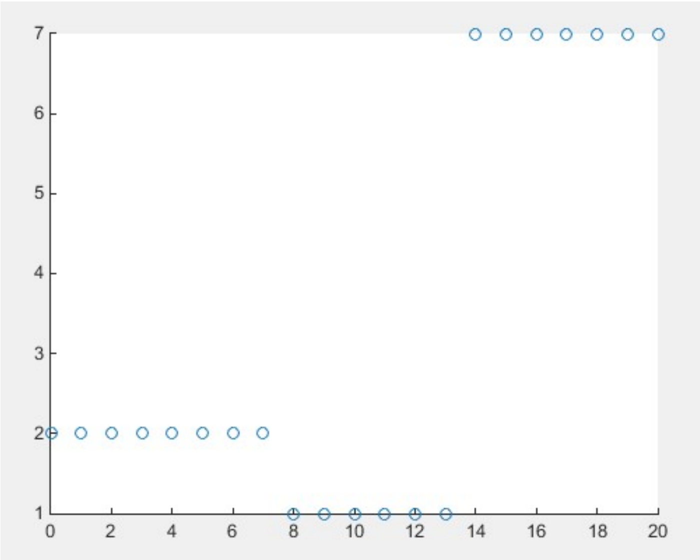
$$+\ln\mathrm{B}\left(\boldsymbol{\alpha}^{(\mathrm{t})}\right)+\left(\sum_{j=1}^{K}\alpha_j^{(t)}-K\right)\psi\left(\sum_{j=1}^{K}\alpha_j^{(t)}\right)-\sum_{j}^{K}\left(\alpha_j^{(t)}-1\right)\psi(\alpha_j^{(t)})$$

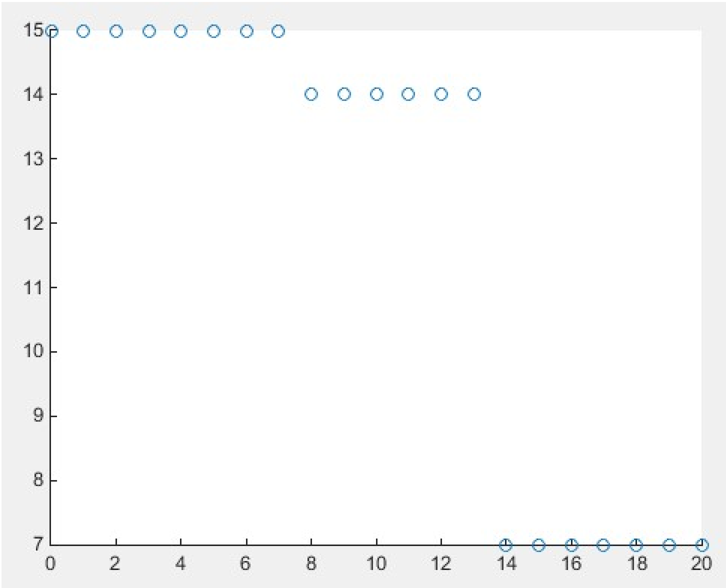
b) Objective function



c) Find most probable clusters







Problem 3

a)

for all j such that $n_j^{(-i)} > 0$

$$p(c_i = j | x_i, \boldsymbol{\theta}, c_{-i}) \propto \frac{p(x_i|\theta_j) n_j^{(-i)}}{\alpha + n - 1}$$

$$\propto \frac{\frac{20!}{x_i!\,(20 - x_i)!} \theta_j^{x_i} (1 - \theta_j)^{20 - x_i} n_j^{(-i)}}{\alpha + n - 1} = \phi_i(j)$$

for a new cluster j'

$$p(c_i = new | x_i, \boldsymbol{\theta}, c_{-i}) \propto \frac{\alpha}{\alpha + n - 1} \int p(x_i|\theta) p(\theta) d\theta$$

$$\propto \frac{\alpha}{\alpha+n-1}\frac{20!}{x_i!\,(20-x_i)!}\frac{B(x_i+a,20-x_i+b)}{B(a,b)} = \phi_i(j')$$

For all cluster

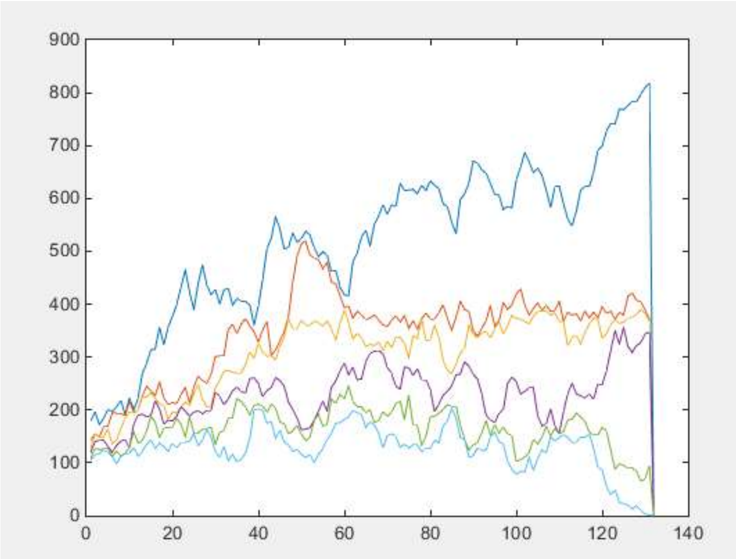
$$p(\theta_j|\{x_i: c_i = j\}) = \frac{\prod p(\{x_i: c_i = j\}|\theta_j)p(\theta_j)}{\int\prod p(\{x_i: c_i = j\}|\theta)p(\theta_j)d\theta_j}$$

$$= \mathrm{Beta}(a', b')$$

$$a' = \sum_{i:c_i=k} x_i + a\,, \qquad b' = \sum_{i:c_i=k} 20 - x_i + b$$

b)

Six most probable clusters



c)

Total number of clusters